



Rat der
Europäischen Union

Brüssel, den 26. Oktober 2022
(OR. en)

14095/22
ADD 1

LIMITE

CORLX 995
CFSP/PESC 1432
CONUN 254
CODUN 48
CONOP 112
COTER 256
COARM 217

VORSCHLAG

Absender: Herr Stefano SANNINO, Generalsekretär, im Auftrag des Hohen Vertreters der Union für Außen- und Sicherheitspolitik

Eingangsdatum: 26. Oktober 2022

Empfänger: Frau Thérèse BLANCHET, Generalsekretärin des Rates der Europäischen Union

Betr.: Vorschlag des Hohen Vertreters der Union für Außen- und Sicherheitspolitik an den Rat für einen Beschluss des Rates zur Unterstützung der Durchführung eines Projekts mit dem Titel „Förderung verantwortungsvoller Innovation auf dem Gebiet der künstlichen Intelligenz im Dienste von Frieden und Sicherheit“

Die Delegationen erhalten in der Anlage das Dokument HR(2022) 238.

Anl.: HR(2022) 238

EUROPÄISCHER AUSWÄRTIGER DIENST



PROJEKTDOKUMENT

„Förderung verantwortungsvoller Innovation auf dem Gebiet der künstlichen Intelligenz im Dienste von Frieden und Sicherheit“

HR(2022)238

1. HINTERGRUND

Durch die jüngsten Fortschritte auf dem Gebiet der künstlichen Intelligenz (im Folgenden „KI“) wurden neue Möglichkeiten geschaffen, Frieden und Sicherheit zu fördern und zu bewahren, beispielsweise durch technologische Verbesserungen in Bereichen wie der Konfliktfrühwarnung und der Verifizierung von Rüstungskontrolle und Ausfuhrkontrolle. Diese Fortschritte haben andererseits auch die Entstehung neuer Mittel begünstigt, mit denen Spannungen, Konflikte und Unsicherheit zwischen und innerhalb von Staaten hervorgerufen – oder verschärft – werden können. Die Risiken, die mit bestimmten KI-Anwendungen wie letalen autonomen Waffensystemen einhergehen, bereiten der Rüstungskontrollgemeinschaft mittlerweile große Sorgen. Es besteht ein Risikopfad, der mehr Aufmerksamkeit verdient und dem mit den gegenwärtigen Rüstungskontrollbemühungen und diplomatischen Bemühungen möglicherweise nicht ausreichend begegnet wird, nämlich die Umlenkung und der Missbrauch ziviler Forschung und Innovation auf dem Gebiet der KI durch verantwortungslose Akteure, einschließlich böswilliger nichtstaatlicher Akteure, beispielsweise der Missbrauch von Generative Adversarial Networks (im Folgenden „GAN“) zur Erstellung von „Deepfakes“ für Desinformationskampagnen.

KI ist eine grundlegende Technologie mit großem Potenzial für die allgemeine Nutzung. Bestimmte Akteure könnten (relativ leicht) auf KI-Forschung und -Innovation, die für zivile Anwendungen bestimmt sind, zugreifen und sie für schädliche oder zerstörerische Zwecke, die sich auf den Weltfrieden und die internationale Sicherheit auswirken könnten, nutzen. Die Umlenkung und der Missbrauch ziviler Technologien sind weder ein neues Phänomen noch auf die KI beschränkt. Auf dem verwandten Gebiet der Robotertechnik wurde dies kürzlich durch die Nutzung von Freizeitdrohnen als Waffe durch Da‘esh/ISIS in Syrien vor Augen geführt. Im Fall der KI ist die Lage aufgrund verschiedener Faktoren komplizierter: KI-Algorithmen und -daten sind immateriell und ändern sich rasch, weshalb ihre Weitergabe/Verbreitung schwer zu kontrollieren ist; der private Sektor spielt im Ökosystem der Forschung, Entwicklung und

Innovation eine führende Rolle, weshalb unternehmenseigene Algorithmen geschützt werden müssen; das Humanwissen und die materiellen Ressourcen, die für die Umnutzung von KI-Technologien benötigt werden, sind weltweit verfügbar. Derweil sind sich diejenigen, die im zivilen Sektor mit KI arbeiten, allzu oft nicht darüber im Klaren, welche Auswirkungen die Umlenkung und der Missbrauch ihrer Arbeit auf den Weltfrieden und die internationale Sicherheit haben könnten, oder sie beteiligen sich nur zögerlich an den Debatten, die in den mit Rüstungskontrolle und Nichtverbreitung befassten Kreisen über die mit KI verbundenen Risiken geführt werden.

Die zivile KI-Gemeinschaft muss dazu gebracht werden, sich stärker einzubringen, wenn es darum geht zu verstehen, welche Risiken für Frieden und Sicherheit die Umlenkung und der Missbrauch ziviler KI-Technologien durch verantwortungslose Akteure mit sich bringen, und diese Risiken zu mindern. Nach Auffassung des Stockholmer Internationalen Friedensforschungsinstituts (im Folgenden „SIPRI“) und des Büros der Vereinten Nationen für Abrüstungsfragen (im Folgenden „UNODA“) könnte dies dadurch erreicht werden, dass in der globalen zivilen KI-Gemeinschaft verstärkt für verantwortungsvolle Innovation geworben wird. Die bisherige Arbeit des SIPRI und des UNODA hat gezeigt, dass verantwortungsvolle Innovation in Form eines Selbstverwaltungsmechanismus der globalen zivilen KI-Gemeinschaft praktische Instrumente und Verfahren an die Hand geben könnte, sodass sie in der Lage wäre, die durch Umlenkung und Missbrauch ziviler KI-Forschung und -Innovation möglicherweise entstehenden Risiken für Frieden und Sicherheit zu erkennen und dazu beizutragen, dass ihnen vorgebeugt wird und sie gemindert werden. Das SIPRI und das UNODA haben im Rahmen ihrer Arbeit zudem Methoden und mehrere laufende auf den zivilen Bereich ausgerichtete Initiativen für verantwortungsvolle KI ermittelt, auf denen aufgebaut werden könnte, um die zivile KI-Gemeinschaft für Abrüstungs- und Nichtverbreitungsfragen, für Expertendebatten und für Standpunkte von Staaten zu verantwortungsvoller Entwicklung, Verbreitung und Nutzung von KI, zu sensibilisieren, und sie haben festgestellt, welche Lehren aus der diesbezüglichen Arbeit im Verteidigungssektor gezogen werden können.¹ Im Zuge dieser früheren Arbeit wurde eindeutig festgestellt, dass die Zusammenarbeit mit Studierenden der Fachbereiche Mathematik, Ingenieurwesen, Naturwissenschaften und Technologie (MINT), noch während diese sich im Rahmen ihrer Ausbildung mit KI befassen, entscheidende Voraussetzung für wirksame Bemühungen um eine verantwortungsvolle Innovation ist.

2. ZIELE

Mit den Projekten soll die zivile KI-Gemeinschaft unterstützt werden, damit sie sich stärker für eine Minderung der Risiken für den Weltfrieden und die internationale Sicherheit einsetzt, die

¹ Zu den Methoden zählen unter anderem die vom Berufsverband der Ingenieure für Elektrotechnik und Informationstechnik (Institute of Electrical and Electronics Engineers - IEEE) empfohlenen Verfahren zur Bewertung der Auswirkung autonomer und intelligenter Systeme auf das menschliche Wohlergehen (IEEE Std 7010-2020), sowie die von der Gruppe hochrangiger Experten für künstliche Intelligenz erstellte Liste zur Bewertung vertrauenswürdiger künstlicher Intelligenz (High-Level Expert Group on Artificial Intelligence' Assessment list for Trustworthy Artificial Intelligence -ALTAI). Zu den Initiativen zählen unter anderem die globale Initiative des IEEE zu ethischen Fragen bei autonomen und intelligenten Systemen (IEEE global initiative on ethics for autonomous and intelligent systems), die Partnerschaft für KI und die Globale Partnerschaft für KI.

von der Umlenkung und vom Missbrauch ziviler KI-Forschung und -Innovation durch verantwortungslose Akteure ausgehen können. Um dies zu erreichen, wird erstens versucht, besser zu verstehen, wie Entscheidungen bezüglich der Entwicklung und der Verbreitung von Forschung und Innovation im Bereich der KI das Risiko der Umlenkung und des Missbrauchs beeinflussen und wiederum Risiken oder Chancen für Frieden und Sicherheit bedeuten können, und zweitens werden verantwortungsvolle Innovationsprozesse, -methoden und -instrumente gefördert, die dazu beitragen können, eine friedliche Anwendung ziviler Innovationen und eine verantwortungsvolle Verbreitung von Kenntnissen auf dem Gebiet der KI zu gewährleisten. Zu diesem Zweck werden im Rahmen der Projekte Maßnahmen zum Aufbau von Kapazitäten, Forschungsmaßnahmen und Maßnahmen zur Einbeziehung unterstützt, durch die i) die globale zivile KI-Gemeinschaft besser befähigt wird, durch verantwortungsvolle Innovationsprozesse die Risiken für Frieden und Sicherheit, die von der Umlenkung und vom Missbrauch ziviler KI durch verantwortungslose Akteure ausgehen, mitzubedenken und zu mindern, und ii) die Bemühungen um Risikominderung durch verantwortungsvolle KI im zivilen Bereich stärker mit den bereits auf zwischenstaatlicher Ebene laufenden Bemühungen in den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung verknüpft werden. Die Projektmaßnahmen haben grundsätzlich nicht zum Ziel, neue Standards, Grundsätze oder Regelungen festzulegen oder anderweitig in Bereiche einzudringen, die in die Zuständigkeit der Staaten fallen. Vielmehr soll erreicht werden, dass im zivilen Bereich verstärkt verantwortungsvolle Innovationsanstrengungen unternommen werden, die den von der Umlenkung und vom Missbrauch ziviler KI durch verantwortungslose Akteure ausgehenden Risiken für Frieden und Sicherheit Rechnung tragen, und dass zu den bestehenden einschlägigen zwischenstaatlichen Bemühungen Schulungen bereitgestellt werden.

Um die zivile KI-Gemeinschaft tatsächlich zu erreichen und auf sie einzuwirken, wird im Rahmen der Projekte ein auf drei Bereiche zielender Ansatz verfolgt:

- (a) Zusammenarbeit mit Pädagogen - Zusammenarbeit mit ausgewählten Pädagogen und Personen, die mit der Aufstellung akademischer Lehrpläne befasst sind, bei der Ausarbeitung und Verbreitung von didaktischem Material, das verwendet werden kann, um zu erreichen, dass die von der Umlenkung und vom Missbrauch ziviler KI-Forschung und -Innovation durch unverantwortliche Akteure ausgehenden Risiken für Frieden und Sicherheit in der Ausbildung der künftigen KI-Praktiker durchgängig berücksichtigt werden (beispielsweise in Kursen zur Ethik in der künstlichen Intelligenz und zu verantwortungsvoller Innovation);
- (b) Einbeziehung von Studierenden - Einführung ausgewählter Studierender der Fachbereiche Mathematik, Ingenieurwesen, Naturwissenschaften und Technologie (MINT) aus aller Welt in die Frage, wie die von der Umlenkung und vom Missbrauch ziviler KI-Entwicklungen durch unverantwortliche Akteure ausgehenden Risiken für Frieden und Sicherheit bei Forschungs- und Innovationsprozessen oder durch andere Governance-Prozesse erkannt werden können, ihnen vorgebeugt werden kann oder sie gemindert werden können, und
- (c) Einbeziehung der Unternehmen der KI-Branche - Zusammenarbeit mit Berufsverbänden und Normungsinstituten wie beispielsweise dem Berufsverband der Ingenieure für Elektrotechnik und Informationstechnik (Institute of Electrical and Electronics Engineers

- IEEE), um i) für Fachleute auf ihren Bedarf abgestimmtes didaktisches Material und auf ihren Bedarf abgestimmte Maßnahmen zur Einbeziehung bereitzustellen, ii) die positive Nutzung von KI für Frieden und Sicherheit zu unterstützen und iii) den Dialog und den Informationsaustausch zwischen Experten aus Wissenschaft und Privatwirtschaft und Regierungsexperten über die Frage zu erleichtern, wie das Risiko der Umlenkung und des Missbrauchs ziviler Forschung und Innovation auf dem Gebiet der KI durch unverantwortliche Akteure gemindert werden kann.

Ein solcher Ansatz sorgt dafür, dass mit den Projekten alle Ebenen der KI-Gemeinschaft erreicht werden, nicht nur die Praktiker, die gegenwärtig auf diesem Gebiet arbeiten, sondern auch künftige Generationen von Praktikern. Er ermöglicht zudem eine Zusammenarbeit von Wissenschaft, Wirtschaft und anderer Sparten und fördert zudem durch den Aufbau spartenübergreifender Netze die Nachhaltigkeit künftiger Anstrengungen.

Die Projekte zielen zudem darauf ab, die Fähigkeit des SIPRI und des UNODA, alle Beteiligten an einen Tisch zu bringen, und ihre diesbezügliche Erfahrung zu nutzen, um auf die AI-Gemeinschaft weltweit, und nicht nur auf die Interessenträger in der EU, einzuwirken. Das SIPRI und das UNODA sind besonders gut in der Lage, KI-Akteure aus Afrika, dem asiatisch-pazifischen Raum, Europa und Nord- und Südamerika zu erreichen und die Zusammenarbeit zwischen ihnen zu begünstigen. Beide Einrichtungen verfügen außerdem über Erfahrung in anderen Bereichen der Wissenschaft und Technologie, in denen im Hinblick auf doppelte Verwendungszwecke und Verbreitung vergleichbare Probleme bestehen, so unter anderem im Bereich der Biotechnologie. Im Rahmen der Projekte sollen ferner die Gegebenheiten in der Europäischen Union genutzt werden, nämlich a) das Vorhandensein fortgeschrittener Multi-Stakeholder-Prozesse zum Thema verantwortungsvolle KI, b) das große Engagement und Fachwissen in Bezug auf Abrüstungs-, Rüstungskontroll- und Nichtverbreitungsfragen in der EU, c) die vielfältigen Verbindungen, die Organisationen aus Wissenschaft, Forschung und Privatwirtschaft in der EU zu anderen Regionen unterhalten, insbesondere im Globalen Süden, der ebenfalls ein wesentliches Ziel der Zusammenarbeit sein wird, und d) die vielen unterschiedlichen Nationalitäten der Studierenden, Pädagogen und Ingenieure, die in Universitäten und Forschungseinrichtungen sowie in der Privatwirtschaft tätig sind.

Der Inklusion wird bei der Durchführung der Projektmaßnahmen besondere Beachtung geschenkt. Um die KI-Gemeinschaft wirksam zu unterstützen, wird im Rahmen der Projekte der Tatsache Rechnung getragen, dass die KI-Gemeinschaft aus einer Vielzahl von Akteuren besteht und dass insbesondere

- (a) das Geschlecht ein äußerst relevanter Faktor ist. Deshalb wird die Gleichstellung der Geschlechter im Einklang mit den im gesamten VN-System verfolgten Gleichstellungs- und Paritätsstrategien durchgängig berücksichtigt. Die Beteiligung von Frauen an allen Projektmaßnahmen wird gefördert und gefordert, und
- (b) die Inklusion von Menschen mit Behinderungen und die Bereitstellung angemessener Vorkehrungen während der gesamten Projektdauer sicherzustellen ist. So werden unter anderem Hindernisse beseitigt, die der Beteiligung von Menschen mit Behinderungen entgegenstehen, sowie Schritte unternommen, um auf Menschen mit Behinderungen

zuzugehen und dafür zu sorgen, dass ihre inhaltlichen Ansichten und Erfahrungen besser berücksichtigt werden.

3. PROJEKTE

Die drei nachstehend beschriebenen Projekte sollen einander ergänzen und unterstützen, wobei sich bestimmte Projektkomponenten über die gesamten 36 Monate der Projektlaufzeit erstrecken werden.

Projekt 1 – Erstellung von didaktischem Material und von Materialien zum Aufbau von Kapazitäten für die zivile KI-Gemeinschaft

3.1. Projektziel

Mit Projekt 1 sollen zivilen Akteuren im KI-Bereich die Kenntnisse und Mittel bereitgestellt werden, die sie in die Lage versetzen, die Risiken für den Weltfrieden und die internationale Sicherheit, die von der Umlenkung und vom Missbrauch ziviler KI-Forschung und -Innovation durch verantwortungslose Akteure ausgehen können, zu evaluieren und zu mindern. Ziel des Projekts ist die Erstellung von didaktischem Material und von Material zum Aufbau von Kapazitäten, das KI-Praktiker aller Regionen, Ebenen und Sektoren (einschließlich Pädagogen, deren Schwerpunkt auf KI liegt, Personen, die mit der Aufstellung von Lehrplänen befasst sind, MINT-Studierender sowie KI-Ingenieure und Forschender der Universitäten und der Privatwirtschaft) mit den Informationen und Instrumenten ausstattet, die sie in die Lage versetzen,

- (a) zu verstehen, auf welche Art und Weise zivile Forschung und Innovation auf dem Gebiet der KI umgelenkt und missbraucht werden kann, sodass daraus Risiken für den Weltfrieden und die internationale Sicherheit entstehen, und wie durch Entscheidungen bezüglich der Entwicklung und Verbreitung von Forschung und Innovation das Umlenkungs- und Missbrauchsrisiko erhöht oder vermindert werden kann,
- (b) sich Einblicke in die Anstrengungen zu verschaffen, die in den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung bereits unternommen worden sind, um das Risiko der Umlenkung und des Missbrauchs von ziviler Forschung und Innovation zu mindern, und
- (c) im Rahmen ihrer innovativen Tätigkeit so verantwortungsvoll vorzugehen, dass das Risiko der Umlenkung und des Missbrauchs bei der Weiterentwicklung und Verbreitung von Forschung und Innovation gemindert wird.

3.1.1. Beschreibung des Projekts

Im Rahmen dieses Projekts werden drei unterschiedliche Gruppen von didaktischem Material und Material zum Aufbau von Kapazitäten erstellt.

- (a) *ein (1) Handbuch* – In dem Handbuch werden die grundlegenden Kenntnisse und Mittel zusammengestellt, die Akteure im KI-Bereich in die Lage versetzen, bei Forschungs- und Innovationsprozessen die Risiken der Umlenkung und des Missbrauchs von ziviler KI-Technologie durch verantwortungslose Akteure zu evaluieren und zu mindern. Ferner

wird darin dargelegt, aus welchen Gründen und auf welche Weise Entscheidungen bezüglich der Entwicklung und Verbreitung von Forschung und Innovation das Umlenkungs- und Missbrauchsrisiko beeinflussen und wiederum Risiken oder Chancen für Frieden und Sicherheit bedeuten können. Außerdem wird das Handbuch neben einer Erläuterung der Sicherheits- und Gefahrenabwehraspekte, die derzeit in militärischen und in mit Abrüstung, Rüstungskontrolle und Nichtverbreitung befassten Kreisen erörtert werden, auch eine Einführung in die einschlägigen völkerrechtlichen Verpflichtungen und Ausfuhrkontrollverpflichtungen sowie Beispiele für Prozesse und Instrumente enthalten, die verantwortungsvolle Innovation ermöglichen, etwa Methoden der Technologiefolgenabschätzung und Vorlagen für die Risikobewertung.

- (b) eine (ca. zehn Beiträge umfassende) Podcast-Reihe – Die Podcasts werden ein zugängliches und ansprechendes Medium sein, das Akteure im KI-Bereich darüber informiert, aus welchen Gründen und auf welche Weise verantwortungsvolle Innovationsprozesse auf dem Gebiet der KI den Weltfrieden und die internationale Sicherheit fördern können, indem sie die Risiken, die von der Umlenkung und vom Missbrauch durch verantwortungslose Akteure ausgehen, mindern. In den Podcasts werden wichtige Themen behandelt (z. B. Umlenkungs- und Missbrauchsmuster im Bereich der Forschung und Innovation mit doppeltem/allgemeinem Verwendungszweck, die humanitären, strategischen und politischen Probleme, die mit dem potenziellen Missbrauch von ziviler Forschung und Innovation auf dem Gebiet der KI einhergehen, die Probleme, mit denen die mit Abrüstung, Rüstungskontrolle und Nichtverbreitung befassten Kreise bei ihren Bemühungen um Risikominderung konfrontiert sind, die Frage, wie Risikobewertung verantwortungsvolle Innovation ermöglichen kann, Compliance im Zusammenhang mit Ausfuhrkontrollen, Risikominderung durch Design, verantwortungsvolle Veröffentlichung, Kenntnis der Kunden und Erfahrung mit Planübungen); sie werden hauptsächlich aus Interviews bestehen, die das Projektteam mit Vertretern der einschlägigen Gemeinschaften führen wird.
- (c) eine (neun bis zehn Beiträge umfassende) Blog-Reihe – Das Team wird eine kuratierte Reihe von Blog-Posts ausarbeiten, die darauf abzielen, die Anstrengungen besser bekannt zu machen, die unternommen werden, um die Grenzen zwischen der auf den zivilen Bereich ausgerichteten „verantwortungsvollen KI“ und den Bereichen Abrüstung und Nichtverbreitung zu überwinden. Die Blog-Reihe wird eine Plattform bieten, über die Erkenntnisse, Ideen und Lösungen bezüglich der Ermittlung und Beseitigung der Risiken, die mit der Umlenkung und dem Missbrauch von KI im zivilen Bereich bei Forschungs- und Innovationsprozessen einhergehen, verbreitet werden können. Der Blog soll die Meinungsvielfalt und die Vielfalt der Blickwinkel im KI-Sektor widerspiegeln.

Die Blogbeiträge werden über die Websites der durchführenden Akteure, ihre Präsenz in den sozialen Medien sowie durch direkte Kommunikation mit den relevanten akademischen Einrichtungen, zivilen Berufsverbänden im KI-Bereich und anderen geeigneten Gruppen veröffentlicht.

3.1.2. Erwartete Projektergebnisse

Im Rahmen des Projekts soll eine neue Reihe von Materialien geschaffen werden, durch die KI-Praktiker des zivilen Bereichs für folgende Fragen sensibilisiert werden können: a) Wie können Forschung und Innovation auf dem Gebiet der zivilen KI so umgelenkt und missbraucht werden, dass daraus Risiken für den Weltfrieden und die internationale Sicherheit erwachsen, b) wie wird diesen Risiken in den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung begegnet und c) wie können KI-Praktiker durch verantwortungsvolle Innovationsprozesse weiter zur Minderung solcher Risiken beitragen?

Hierdurch soll der zivile KI-Sektor dazu gebracht werden, sich stärker für eine Minderung der von der Umlenkung und vom Missbrauch ziviler KI ausgehenden Risiken für den Weltfrieden und die internationale Sicherheit einzusetzen. Ferner sollen die Praktiker im technischen Bereich besser befähigt werden, eine Zusammenarbeit mit den einschlägigen Prozessen in den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung aufzunehmen und auf die Einbeziehung neuer Zielgruppen hinzuwirken, die traditionell nicht in die didaktischen Bemühungen in den Bereichen Abrüstung und Nichtverbreitung einbezogen sind.

Ferner soll das Material bei der Durchführung anderer Projekte helfen und als Grundlage für die didaktischen Maßnahmen und die Maßnahmen zum Aufbau von Kapazitäten im Rahmen von Projekt 2 sowie für die Dialog- und Einbeziehungsmaßnahmen des Projekts 3 dienen. Diese Maßnahmen sollen wiederum in die Ausarbeitung und Verbesserung des Materials einfließen. Dieser iterative Ansatz soll dabei helfen, potenzielle Hindernisse auszuräumen, die der Propagierung, Verbreitung und Nutzung der Materialien in der KI-Gemeinschaft entgegenstehen, was Probleme in Bezug auf Sprache, Inhalt, Kontext und Verfügbarkeit einschließt, die auf globaler Ebene, und insbesondere im Globalen Süden, die Projektwirkung negativ beeinträchtigen könnten.

3.2. Projekt 2 – Didaktische Maßnahmen und Maßnahmen zum Aufbau von Kapazitäten für künftige KI-Praktiker

3.2.1. Projektziel

Projekt 2 hat zum Ziel, das Problem der Umlenkung und des Missbrauchs von ziviler Forschung auf dem Gebiet der KI durch verantwortungslose Akteure bei der Ausbildung künftiger Generationen von KI-Praktikern einzubeziehen. Auf lange Sicht wird hierdurch sichergestellt, dass Studierende der Bereiche Mathematik, Ingenieurwesen, Naturwissenschaften und Technologie (MINT), die die Zukunft der KI gestalten werden, sich der negativen Auswirkungen, die die Umlenkung und der Missbrauch ihrer Arbeit durch verantwortungslose Akteure auf den Weltfrieden und die internationale Sicherheit haben können, bewusst sind, und dass den Studierenden die erforderlichen grundlegenden Instrumente zur Verfügung stehen, um diese Risiken bei den Forschungs- und Innovationsprozessen zu erkennen und zu mindern.

Im Rahmen des Projekts wird in Zusammenarbeit mit ausgewählten internationalen Universitäten und Akteuren der Wirtschaft eine Reihe von didaktischen Workshops und Workshops zum Aufbau von Kapazitäten durchgeführt, die sich an Pädagogen und Studierende richten. Damit wird das Ziel verfolgt, Maßnahmen zum Aufbau von Kapazitäten zu entwickeln,

die von Pädagogen und Personen, die mit der Aufstellung akademischer Lehrpläne befasst sind, genutzt werden könnten, um bei der Ausbildung künftiger KI-Praktiker (z. B. in Kursen zu Ethik im KI-Bereich und zu verantwortungsvoller Innovation) auch die Risiken zu berücksichtigen, die von der Umlenkung und vom Missbrauch ziviler KI-Forschung und -Innovation durch verantwortungslose Akteure ausgehen, und diese Frage im größeren Kontext von Frieden und Sicherheit zu behandeln. Mit dem Projekt wird zudem das Ziel verfolgt, im Zuge der Workshops ein Netz interessierter Pädagogen, mit der Lehrplanaufstellung befasster Personen und Studierender zu bilden, die sich dafür einsetzen, das didaktische Material und die Maßnahmen zum Aufbau von Kapazitäten in der KI-Bildungsgemeinschaft und der Gemeinschaft der KI-Praktiker zu verbreiten und zu propagieren. Diese Vernetzungskomponente soll dafür sorgen, dass die Projekte über ihre unmittelbare Laufzeit hinaus Wirkung entfalten, und den Aufbau stärkerer Verbindungen ermöglichen, die das Engagement ziviler technischer Kreise für die umfassenderen Ziele Frieden, Sicherheit, Abrüstung und Rüstungskontrolle fördern.

3.2.2. Beschreibung des Projekts

Im Rahmen des Projektes wird eine Reihe von didaktischen Workshops und Workshops zum Aufbau von Kapazitäten durchgeführt, die sich an Pädagogen und Studierende ausgewählter Universitäten aus aller Welt richten. Die Workshops sollen aus einer Mischung von Vorlesungen und interaktiven Angeboten bestehen, die Pädagogen und Studierenden die Möglichkeit geben, Überlegungen zu der Frage anzustellen, wie zivile KI-Forschung und -Innovation so umgelenkt und missbraucht werden könnte, dass Risiken für den Weltfrieden und die internationale Sicherheit entstehen, und wie diese Risiken bei Forschungs- und Innovationsprozessen oder durch andere Governance-Prozesse erkannt werden können, ihnen vorgebeugt werden kann oder sie gemindert werden können. Die Projektmaßnahmen werden auf der vom UNODA vorab in kleinerem Maßstab durchgeführten Pilotarbeit aufbauen, bei der Methoden erprobt wurden, um mit MINT-Studierenden zusammenzuarbeiten und sie dafür zu sensibilisieren, wie wichtig es ist, die Auswirkungen ihrer Arbeit zu bedenken und mit Experten außerhalb des angestammten Fachbereichs zusammenzuarbeiten. Konkret würde es sich um Folgendes handeln:

- (a) *vier (4) an Pädagogen und Studierende gerichtete regionale Workshops zum Aufbau von Kapazitäten* – Im Rahmen der regionalen Workshops werden Maßnahmen durchgeführt und propagiert, die von Pädagogen genutzt werden können, um die Fähigkeiten der MINT-Studierenden in Bezug auf verantwortungsvolle Innovation auf dem Gebiet der KI zu verbessern, wobei der Schwerpunkt auf der Frage liegen soll, wie sich das Risiko, dass zivile KI-Technologie durch verantwortungslose Akteure umgelenkt oder missbraucht wird, evaluieren und mindern lässt. Jeder Workshop ist so konzipiert, dass eine Hochschule aus der EU mit einer profilierten Hochschule aus einer anderen globalen Region zusammentrifft, sodass stets Verbindungen zwischen unterschiedlichsten Teilnehmenden aus der EU mit unterschiedlichsten Teilnehmenden von außerhalb der EU entstehen. Die Workshops werden Lateinamerika und den karibischen Raum, Nordamerika, Afrika und den asiatisch-pazifischen Raum abdecken. Somit werden Studierende (auf Master- und Promotionsebene) aus der ganzen Welt, auch aus dem Globalen Süden, teilnehmen können. Die Workshops sollen in erster Linie in englischer Sprache durchgeführt werden; soweit möglich, werden die Teilnehmenden auch die

Möglichkeit erhalten, Aktivitäten beizuwohnen, die auf andere Sprachgruppen ausgerichtet sind; und

- (b) *ein (1) internationaler Workshop zum Aufbau nachhaltiger Kapazitäten* – Der Workshop wird auf den bei den regionalen Workshops gewonnenen Erfahrungen aufbauen und den Austausch von Informationen und Erfahrungen zwischen Pädagogen und ausgewählten Studierenden der am Projekt teilnehmenden Universitäten erleichtern. Bei diesem Workshop wird erörtert, wie die im Zuge des Projekts erarbeiteten Maßnahmen und Instrumente verbessert und über die Gruppe der teilnehmenden Universitäten hinaus verbreitet werden können. Ferner wird dabei erörtert, wie Studierende dazu gebracht werden können, sich für verantwortungsvolles Handeln auf dem Gebiet der KI, das den von der Umlenkung und vom Missbrauch ausgehenden Risiken für den Weltfrieden und die internationale Sicherheit Rechnung trägt, einzusetzen, sobald sie in das Erwerbsleben eingetreten sind.

Die Netze und die Präsenz des SIPRI und des UNODA in Afrika, im asiatisch-pazifischen Raum, in Europa sowie in Nord- und Südamerika werden genutzt, um Komponenten der Maßnahmen in geeigneter Weise zu erleichtern und zu unterstützen.

3.2.3. Erwartete Projektergebnisse

Im Rahmen des Projektes sollen Modelle für Maßnahmen zum Aufbau von Kapazitäten und Maßnahmen zur Einbeziehung geschaffen werden, die Pädagogen und mit der Aufstellung akademischer Lehrplänen befasste Personen nachbilden können, um künftige KI-Praktiker für das Problem der Umlenkung und des Missbrauchs von ziviler KI durch verantwortungslose Akteure sowie für die Frage zu sensibilisieren, wie sie dazu beitragen können, dieses Problem durch verantwortungsvolle Innovationsprozesse zu mindern. Nach Abschluss der Projektmaßnahmen sollen die Teilnehmenden (Pädagogen, aber auch MINT-Studierende) in der Lage sein, verantwortungsvolle Instrumente, Methoden und Konzepte der Innovation zu nutzen und zu propagieren, die es erlauben, die Umlenkungs- und Missbrauchsrisiken bei der Entwicklung und Verbreitung von ziviler Forschung und Innovation auf dem Gebiet der KI zu erkennen und zu mindern.

Außerdem soll im Zuge der Projektmaßnahmen ein Netz von Pädagogen, mit der Lehrplanaufstellung befassten Personen und Studierenden entstehen, die nicht nur die Projektmaßnahmen innerhalb der KI-Bildungskreise und der KI-Fachkreise (beispielsweise auf Konferenzen der IEEE Computational Intelligence Society) besser bekannt machen, sondern auch bereit sind, mit ihren technischen Kapazitäten zu unter der Ägide der Staaten durchgeführten internationalen Governance-Prozessen (z. B. zum Prozess im Rahmen des Übereinkommens über bestimmte konventionelle Waffen zu neu entstehenden Technologien auf dem Gebiet der letalen autonomen Waffensysteme) beizutragen.

Der kurz- und langfristige Wert dieser Maßnahmen wird durch vor und nach den Maßnahmen durchgeführte Erhebungen nachgewiesen.

3.3. Projekt 3 – Förderung einer längerfristig nachhaltigen Entwicklung, Verbreitung und Wirkung von verantwortungsvoller Innovation auf dem Gebiet der KI im Dienste von Frieden und Sicherheit

3.3.1. Projektziel

Mit Projekt 3 wird das Ziel verfolgt, eine längerfristig nachhaltige Entwicklung, Verbreitung und Wirkung von verantwortungsvoller Innovation auf dem Gebiet der KI zu fördern, um die durch Umlenkung und Missbrauch ziviler KI-Forschung und -Innovation entstehenden Risiken für Frieden und Sicherheit zu mindern. Dies soll durch Rundtischgespräche mit Unternehmen der KI-Branche, Dialoge zwischen verschiedenen Interessenträgern, Erstellung eines öffentlichen Berichts und gezielte Verbreitungsmaßnahmen geschehen. Durch das Projekt soll sichergestellt werden, dass die geleistete Arbeit, insbesondere die Ausbildungsmaßnahmen, die Maßnahmen zum Aufbau von Kapazitäten und die Maßnahmen zur Einbeziehung, die KI-Gemeinschaft insgesamt auf allen Ebenen (angefangen bei Studierenden bis hin zu Ingenieuren und anderen KI-Fachleuten) und über alle geografischen, sektoralen und anderen Grenzen hinweg erreicht und beeinflusst. Um eine möglichst umfassende und tiefgreifende Wirkung zu erzielen, ist es von entscheidender Bedeutung, mit Berufsverbänden in diesem Bereich, wie beispielsweise dem IEEE, zusammenzuarbeiten und mehrdimensional, unter Einbeziehung der Wissenschaft, der Wirtschaft und anderer Sparten, vorzugehen. Auf diese Weise erhalten interessierte Vertreter der verschiedenen KI-Gemeinschaften die Möglichkeit, Verantwortung für das Problem zu übernehmen und darzulegen, wie aus ihrer Sicht Bemühungen um Risikominderung in der gesamten globalen KI-Gemeinschaft durchgeführt und nachhaltig gefördert werden können. Damit das Projekt für Staaten, zwischenstaatliche Organisationen und andere Akteure dauerhaften Wert hat, ist es außerdem wichtig, dass KI-Praktiker von Regierungsexperten, die sich mit der Risikominderung in den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung befassen, lernen und mit ihnen zusammenarbeiten können. Für die Nachhaltigkeit ist es überdies von entscheidender Bedeutung, dass die im Zuge der Maßnahmen zur Einbeziehung gewonnenen Erkenntnisse in geeigneter Weise analysiert, konsolidiert und verbreitet werden.

3.3.2. Beschreibung des Projekts

Das Projekt umfasst die folgenden Haupthandlungsbereiche:

- (a) bis zu neun (9) Dialoge zwischen verschiedenen Interessenträgern zum Thema „Verantwortungsvolle Innovation auf dem Gebiet der KI im Dienste von Frieden und Sicherheit“ – Bei dieser Reihe von Dialogtreffen, die online stattfinden werden, sollen Experten aus Wissenschaft, Forschung, Privatwirtschaft und aus dem Bereich der traditionellen Rüstungskontrolle aus der EU und anderen Teilen der Welt zusammenkommen, um Folgendes zu erörtern:
 - i. technologische Entwicklungen, aus denen sich Verbreitungs-, Umlenkungs- oder Missbrauchsrisiken mit Auswirkungen auf den Weltfrieden und die internationale Sicherheit ergeben können,

- ii. die Frage, wie durch verantwortungsvolle Innovationsprozesse, -verfahren und -mittel zur Risikominderung beigetragen werden kann, sowie Chancen und Herausforderungen für einen Dialog und Wissensaustausch zwischen Interessensgruppen, einschließlich solcher, die in anderen Branchen, beispielsweise im Biotechnologie- und im Chemiesektor, tätig sind, und
- iii. den potenziellen Wert und Zweck sowie das potenzielle Format eines autonomen Expertennetzes und von Dialogen. Die Expertengruppe wird mehrmals im Jahr zusammenkommen und zwei öffentliche Veranstaltungen für die Gemeinschaft insgesamt vorbereiten.

Zwei der neun Online-Treffen sollen für die Öffentlichkeit zugänglich sein, um eine umfassendere Konsultation zu ermöglichen.

(b) bis zu sechs (6) Rundtischgespräche mit Teilnehmenden aus der Privatwirtschaft – Diese Reihe von Rundtischgesprächen, die online stattfinden werden, soll dazu dienen, mit Akteuren der Privatwirtschaft, die mit verantwortungsvollen Innovationsprozessen auf dem Gebiet der KI (z. B. Partnerschaft für KI) befasst sind, einen Dialog über die Frage einzuleiten, wie sie dazu beitragen können, das Risiko, dass zivile KI-Technologien durch verantwortungslose Akteure umgeleitet und missbraucht werden, auf ein Mindestmaß zu reduzieren; zudem soll ausgelotet werden, welche Anreize im Rahmen der Entwicklung der Privatwirtschaft hierfür geschaffen werden könnten. Folgende Themen werden behandelt:

- i. Die Bedeutung des internationalen Sicherheits- und Abrüstungskontextes für die Privatwirtschaft,
- ii. die rechtlichen Rahmenbedingungen für Entwicklung, Einsatz und Betrieb von KI weltweit,
- iii. die Frage, wie an Mechanismen zur Risikobewertung und andere Maßnahmen, die Teil von verantwortungsvollen Innovationsprozessen und von Compliance-Programmen von Unternehmen sind oder werden könnten, angeknüpft werden kann, und
- iv. Erkenntnisse, die in anderen Branchen, Prozessen und Rahmen mit Bezug zur Rüstungskontrolle (z. B. im Biotechnologie- und im Chemiesektor) gewonnen wurden.

(c) ein (1) an die KI-Gemeinschaft und den Bereich Rüstungskontrolle gerichteter Bericht über die Perspektiven der KI-Gemeinschaft in Bezug auf Rüstungskontrolle und Risikominderung auf dem Gebiet der KI – Im Zuge der Ausarbeitung dieses Bericht werden die Projektergebnisse und die im Rahmen des Projekts ausgesprochenen Empfehlungen in einem einzigen Referenzdokument erfasst und konsolidiert; Zielgruppen sind die auf den zivilen Bereich ausgerichtete verantwortungsvolle KI-Gemeinschaft sowie der Rüstungskontrollsektor. In dem Bericht wird dargelegt, wie die mit der Umlenkung und dem Missbrauch ziviler KI-Forschung und -Innovation verbundenen Risiken für den Weltfrieden und die internationale Sicherheit erkannt, bewertet und ausgeräumt werden können.

(d) Informationsveranstaltungen zur Konsultation und Einbindung der KI-Gemeinschaft und des Rüstungskontrollsektors (Anzahl noch festzulegen) – Das Team wird versuchen, während der gesamten Projektlaufzeit über die Arbeit, die Arbeitsergebnisse und die sich daraus ergebenden Empfehlungen zu informieren. Das Veranstaltungsformat und der Inhalt der Präsentationen werden auf den Bedarf der einzelnen Zielgruppen zugeschnitten. Hierzu können Sitzungen der Gruppe „Nichtverbreitung und Waffenausfuhren“, der „European AI Alliance Assembly“, der Gruppe der Regierungsexperten für neu entstehende Technologien im Bereich der letalen autonomen Waffensysteme, der interinstitutionellen Arbeitsgruppe zu KI (IAWG-AI), der ITU-Initiative „AI for Good“, des „Annual Innovation Dialogue“ (jährlicher Innovationsdialog) des Instituts der Vereinten Nationen für Abrüstungsforschung (UNIDIR) sowie des IEEE gehören. Das Team wird sich ebenfalls bemühen, bilateral mit den relevanten Interessenträgern aus Regierungs-, Wissenschafts- und Privatwirtschaftskreisen zusammenzuarbeiten.

3.3.3. Erwartete Projektergebnisse

Mit dem Projekt soll die Grundlage für eine nachhaltige Entwicklung, Verbreitung und Wirkung von verantwortungsvoller Innovation bei KI-Prozessen gelegt werden, damit die Verbreitungs-, Umlenkungs- und Missbrauchsrisiken und die Folgen dieser Risiken für Frieden und Sicherheit auch über die unmittelbare Geltungsdauer des Ratsbeschlusses hinaus thematisiert werden.

Der Dialog zwischen verschiedenen Interessenträgern soll als Modell für den Informationsaustausch und die Zusammenarbeit in Bezug auf die Risikominderung nicht nur innerhalb der globalen KI-Gemeinschaft, sondern auch zwischen der auf den zivilen Bereich ausgerichteten verantwortungsvollen KI-Gemeinschaft und den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung dienen. Dieses Modell könnte genutzt werden, um politische Entscheidungsträger mit den wichtigsten technologischen und wissenschaftlichen Fortschritten, die für eine verantwortungsvolle Innovation auf dem Gebiet der KI relevant sind, vertraut zu machen, und zudem technischen Kreisen näher zu bringen, in welchem Umfeld die politischen Entscheidungsträger gegenwärtig tätig sind. Das Projekt soll nachhaltige Beziehungen und eine dauerhafte Zusammenarbeit zwischen interessierten Akteuren innerhalb der verschiedenen Gemeinschaften und zwischen den verschiedenen Gemeinschaften erleichtern. Die entsprechenden Effekte heterogener Netze dürften eine stärkere Entwicklung und eine breite Förderung von verantwortlicher Innovation auf dem Gebiet der KI im Dienste von Frieden und Sicherheit auch über den Zeitrahmen des Projekts hinaus ermöglichen.

Der Dialog mit der Privatwirtschaft dürfte bewirken, dass sich die auf dem Gebiet der KI tätigen Unternehmen der Privatwirtschaft stärker und intensiver mit der Frage befassen, wie die durch Umlenkung und Missbrauch ziviler KI-Forschung und -Innovation entstehenden Risiken für Frieden und Sicherheit erkannt werden können, ihnen vorgebeugt werden kann und sie gemindert werden können. Das Projekt soll bewirken, dass die Hauptakteure der privatwirtschaftlichen Prozesse ein besseres Verständnis und mehr Eigenverantwortung für die Problematik entwickeln. Außerdem soll es dazu beitragen, dass im Rahmen bestehender

Risikomanagementsysteme und -verfahren der Unternehmen (verstärkt) verantwortungsvolle Innovationsprozesse, -verfahren und -mittel übernommen und angewandt werden.

Der Dialog zwischen verschiedenen Interessenträgern und das Rundtischgespräch mit Teilnehmenden aus der Privatwirtschaft sollen zudem zu Erkenntnissen in Bezug auf einige wesentliche Fragen führen, unter anderem in Bezug auf folgende Fragen: a) Wie können verantwortungsvolle Methoden und Mittel der Innovation weiter verbessert und eingesetzt werden, um die mit der Umleitung und dem Missbrauch ziviler KI-Forschung und -Innovation einhergehenden Risiken zu erkennen, ihnen vorzubeugen und sie zu mindern? b) Wie können Forschung und Innovation auf dem Gebiet der KI positiv für die Verwirklichung von Friedens- und Sicherheitszielen genutzt werden (z. B. Anwendungen für die Konfliktfrühwarnung und humanitäre Hilfe)? c) Welche Möglichkeiten bestehen, um einen intensiveren Dialog und Informationsaustausch zwischen den Risikominderungsbemühungen der auf den zivilen Bereich ausgerichteten verantwortungsvollen KI-Gemeinschaft (beispielsweise den verschiedenen Initiativen unter der Federführung des IEEE) und den bereits auf zwischenstaatlicher Ebene laufenden Bemühungen in den Bereichen Abrüstung, Rüstungskontrolle und Nichtverbreitung zu erleichtern?

Im Zuge der Berichterstellung und der Maßnahmen zur Informationsverbreitung werden die durch die Projekte 1, 2 und 3 gewonnenen Erkenntnisse analysiert, konsolidiert und verbreitet, um die Projektergebnisse und daraus resultierenden Empfehlungen sowohl in der globalen KI-Gemeinschaft als auch in Kreisen der Politik besser bekannt zu machen. Dies soll auch dazu beitragen, dass die Projekte über ihre Laufzeit hinaus Wirkung entfalten.

4. LAUFZEIT

Die Dauer der Durchführung der Projekte wird mit insgesamt 36 Monaten veranschlagt.